

JAPANESE

[JP,10-040260,A]

CLAIMS DETAILED DESCRIPTION TECHNICAL  
FIELD PRIOR ART EFFECT OF THE INVENTION  
TECHNICAL PROBLEM MEANS DESCRIPTION OF  
DRAWINGS DRAWINGS

[Translation done.]

## \* NOTICES \*

JPO and INPIT are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

## DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Field of the Invention] This invention uses the text in an image as an index, and relates to the image retrieval method of showing the section of the image corresponding to the context by the side of the user inputted as retrieval required.

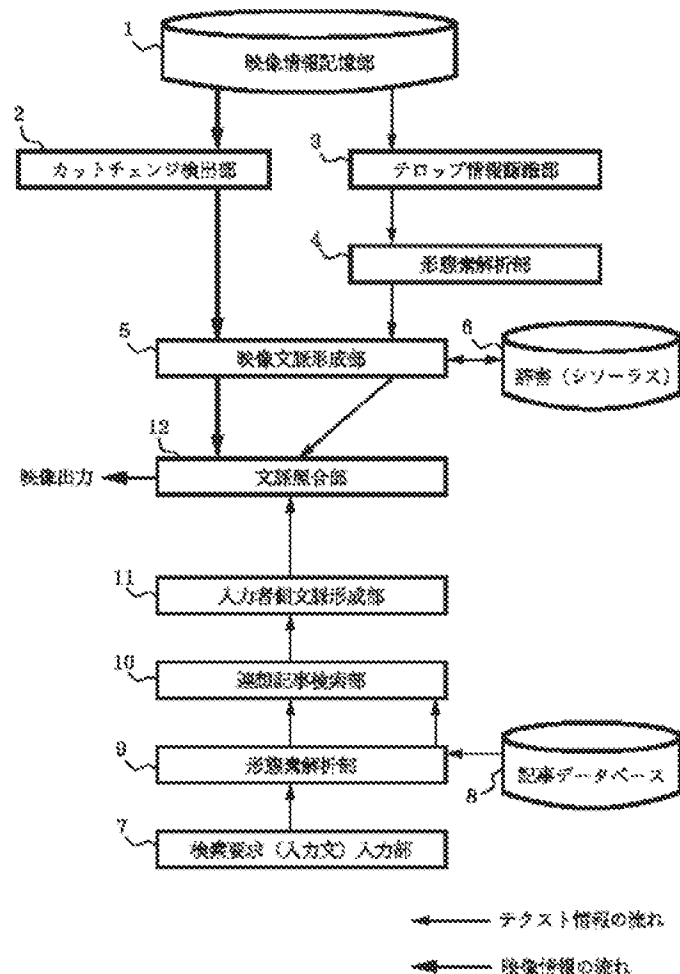
[0002]

[Description of the Prior Art] The technique of recent years many is proposed about the information retrieval technique for video information, or the video information structure technique aiming at search. Also in it, the art (references 1 and 2) of aiming at realization of image retrieval is proposed by inputting the semantic content description about an image as retrieval required.

[0003] The references 1 and 2 are the techniques of structurizing an image using the text information (scenario information is used and, as for the reference 2, the reference 1 uses caption data) added to the image, and searching the

Drawing selection

Representative draw



[Translation done.]

arbitrary image sections by semantic description.

[0004]Reference:[1] Takeshita Atsushi, "recognition of the subject using the interaction structure of the dialog", the Information Processing Society of Japan research technical report 87-10, pp.75-82 (1992).

[0005][2] Yoshitomo Yaginuma, Masao Sakauchi, one proposal of how to match the drama image, sound, and scenario document using DP matching, IEICE TRANSACTIONS, Vol. J79-D-II, No.5, and pp.747-755 (1996).

[0006]

[Problem(s) to be Solved by the Invention]The references 1 and 2 are premised on the case where the text information which synchronized with the image in contents like the title method or the scenario is acquired, and are the techniques of structurizing the image which synchronized with the text in the analysis of an interactive subject structure of text information. However, when aimed at a general broadcast image etc., generally it is difficult [ it ] for text information with the contents which synchronized with image contents, such as title broadcast and scenario information, semantically to come to hand in many cases.

[0007]The purpose of this invention is to provide the image retrieval method of showing the section of the image corresponding to a user's context inputted as retrieval required with the technique on condition of title broadcast or existence of a scenario.

[0008]

[Means for Solving the Problem]A cut change detection stage in which an image retrieval method of this invention detects a point changing [ cut ] from a frame image sequence in an image, A telop information recognition stage which starts a telop out of an image and is recognized as text information, A morphological-analysis stage of applying recognized text information to a morphological analysis, and extracting a concept, An image context formation stage which extends an extracted concept to a generic concept using a dictionary, and gives an extended generic concept as an index to the macroscopic section which consists of two or more cuts in an image, A retrieval-required input stage story into which a user inputs retrieval required as an input sentence, and a morphological-analysis stage of applying an input sentence inputted by user to a morphological analysis, and extracting a concept, An associative report search stage which develops an extracted conceptual group associationally using reports, such as a newspaper, a magazine, etc. accumulated independently of an image, The

input person side context formation stage which takes a set union of a conceptual group extracted in said morphological-analysis stage and said associative report search stage, and forms the context by the side of an input person. It has a context collation step which presents the image section which the context of an image which matched the context by the side of an input person and the context of an image acquired in said image context formation stage, and was matched with the context by the side of an input person expresses as an image of search results.

[0009]This invention is used as an index for search of comparatively general telop information as text information added in an image. It is a method for realizing search of an image by semantic description, and uses that telop information is given a cut which is the semantic minimum unit of an image as the feature by unit.

[0010]First, a point of an image changing [ cut ] is automatically detected using the technique of the references 5 and 6. Next, start a telop out of an image using a method of the references 3 and 4, and it recognizes as text information. Next, a concept which applies text information recognized from a telop to a morphological analysis, and considers it as an index to a cut which extracted a concept (a vocabulary, especially a noun), and from which a telop was detected, next is acquired from a telop is received. Investigate using a dictionary (thesaurus) whether a generic concept shared with a concept from a telop obtained next on a time-axis exists, and in existing. When both concepts are acquired, the context of an image is formed by giving a generic concept as an index to the macroscopic section (the section in which a telop is shown and which is not is also included) to which the section is connected.

[0011]Next, apply an input sentence equivalent to retrieval required inputted by user to a morphological analysis, and a conceptual group (vocabulary group) is extracted. Next, out of article databases, such as a newspaper, a magazine, etc. accumulated independently of an image. The context by the side of an input person is formed by adding a conceptual group extracted from an associative report by conceptual group which searched a report which shares an input sentence and most many vocabularies as an associative report, extracted a conceptual group (vocabulary group) also from an associative report by a morphological analysis, next was extracted from an input sentence.

[0012]The image section which the context of an image matched with the context by the side of an input person expresses is shown as an image of search results by

searching the image section which finally has the context which is in agreement with a vocabulary in the context by the side of an input person one after another, and showing it. [0013]Telop information is common as compared with title broadcast or scenario information as text information which synchronized also in contents also in time to video information, and can generalize an object image of search by using a telop as an index for image retrieval as compared with a conventional method.

[0014]however -- as a fault when using telop information as an index - it is necessarily given to no cuts, but there is the feature that very little amount of information as - text information is as compared with title broadcast or scenario information. With therefore, the associational input person side context formation method using article information which an image context formation method which used a dictionary (thesaurus), and an image accumulated independently in this method. Shortage of text information is compensated and it is coped with by realization of a search method using an indirect correspondence relation by a concept of the context.

[0015]Therefore, this method is the art of editing and showing the image section contextually related to an input rather than art of searching the specific section in an image in target search, and the section which does not necessarily have a direct correspondence relation may be shown.

However, to a user without a clear intention like a target, rather, an indirect correspondence relation accompanied by a loose interpretation like this method is considered to become effective sources of information, and can apply to support of a consulting service, support of thinking, etc.

[0016]Reference: It is the 1994 \*\*\*\*\* spring convention D-427 (1994) about search of a library picture by recognition of [3] Atsushi Nemoto, Seiichiro Hanya, Kazuhiro Miyauchi, and a telop.

[0017][4] examination of key target indexing in an image using Shoji Kurakake, Hidekatsu Kuwano, Hiroyuki Arai, Kazumi Odaka, and recognition art, Shingaku Giho IE95-150, PRU95-237, and pp.15-20 (1996).

[0018][5] Miyatake Kyo, Yoshizawa An automatic detection method of a cut which paid its attention to a rate of change of Satoshi, the Ueda \*\*\*\*\* and a frame correlation coefficient, 1990 \*\*\*\*\* spring convention D-299 (1990).

[0019][6] Video browsing, Shingaku Giho IE90-6 using Seita Otsuji, Yoshinobu Tomura, Yuji Oba, and brightness information (1990).

[0020]

[Embodiment of the Invention]Next, an embodiment of the invention is described with reference to drawings.

[0021]Drawing 1 is a block diagram showing the composition of the device which enforces the image retrieval method by the embodiment of this invention.

[0022]The video information storage parts store 1 in which this device accumulates the video information which is a retrieval object, The cut change primary detecting element 2 which detects the point changing [ cut ] automatically from the frame image sequence in an image, The telop information recognition part 3 which starts a telop out of an image and is recognized as text information, The morphological-analysis part 4 which extracts a concept from the text information recognized from the telop by a morphological analysis, It is investigated using the dictionary (thesaurus) 6 whether the generic concept shared with the concept from the telop obtained next on a time-axis exists to the dictionary (thesaurus) 6 and the concept acquired from a telop, By giving a generic concept as an index to the macroscopic section (the section when the telop is not shown is also included) to which the section is connected when both concepts are acquired, in existing, The image context formation part 5 which forms the context of an image, and the retrieval request input part 7 for a user to input as an input sentence which carries out retrieval required, The associative report retrieval part 10 which searches the report which shares an input sentence and most many vocabularies as an associative report out of the article database 8 with the article database 8 which accumulates the report of a newspaper and a magazine, and the morphological-analysis part 9 which extracts a conceptual group from an input sentence by a morphological analysis, The input person side context formation part 11 which forms the context by the side of an input person by adding the conceptual group extracted from the associative report by the conceptual group extracted from the input sentence, By searching the image section with the context which is in agreement with the vocabulary in the context by the side of an input person one after another, and showing it, it comprises the context collating part 12 which presents the image section which the context of the image matched with the context by the side of an input person expresses as an image of search results.

[0023]Drawing 2 is a figure showing the example of processing of the device of drawing 1.

[0024]The cut change primary detecting element 2 detects the point changing [ cut ] from the image accumulated in the

video information Records Department 1, and it is an image 1-10 ... It divides into a cut.

[0025]In the video information storage parts store 1, it is 1-10... The telop which includes concepts, such as a "Hall of the Great Buddha", "Nigatsu-do", "Sangatsu-do", "Sarusawano-ike", "Kofuku-ji", a "five-storied pagoda", and the "south \*\*\*\*\*", to the picture sequence which consists of cuts shall be displayed. These telop information is recognized as text information using the telop information recognition part 3 and the morphological-analysis part 4, and a concept (a "Hall of the Great Buddha", "Nigatsu-do", "Sangatsu-do", "Sarusawano-ike", etc.) is extracted.

[0026]Next, in the video information storage parts store 1, it is investigated using a dictionary (thesaurus) whether the generic concept to share exists between the concept "Hall of the Great Buddha" extracted from the cut 1 on the time-axis of an image, and the concept "Nigatsu-do" extracted from the cut 3 which appears in the next. Since the concept of "Todaiji" is shared in the case of this example, in the case of the macroscopic section to which the cut from which both concepts were acquired is connected, and this example, a generic concept "Todaiji" is given as an index to the cuts 1-3. In the image context formation part 5, such processing is performed to all the cuts, and the context (what extracted the macroscopic section and generic concept to the video information of a yuan) of an image is formed.

[0027]In the example of drawing 2, the input sentence "after seeing Great Buddha of Nara, visit Sarusawano-ike" equivalent to retrieval required should be inputted by the user from the retrieval request input part 7. The morphological-analysis part 9 extracts a conceptual group ("Nara", "Great Buddha", "Sarusawano-ike") from an input sentence.

[0028]Next, by the associative report retrieval part 10, the report which shares an input sentence and most many concepts out of the article database 8 is searched as an associative report, and the searched report also performs morphological-analysis processing and extracts a concept. It should be searched [ the report which shares an input sentence and two concepts ("Nara", "Great Buddha"), in the case of this example, "a souvenir shop being located in a line with the approach in front of Todaiji and a Hall of the Great Buddha, and the deer rice cracker also being sold, if it is called the sightseeing spot of Nara", and ]. Furthermore, morphological-analysis processing is performed to this report, and a concept is extracted.

[0029]Next, the set union of the conceptual group extracted

from the input sentence and the report is taken by the input person context formation part 11, The context by the side of an input person and ""Nara", a "sightseeing spot", "sightseeing", a "spot", "Todaiji", a "Hall of the Great Buddha", "Great Buddha", "before", an "approach", a "souvenir shop", a "souvenir thing", a "souvenir", a "deer rice cracker", a "deer", a "rice cracker", and "Sarusawano-ike"" are formed.

[0030]Finally, by the context collating part 12, the context of an image and the context by the side of an input person are compared, and, in the case of the context by the side of an input person, the context of an image with a common concept, and this example, the cuts 1-3 and the cut 5 are shown as an image of search results.

[0031]

[Effect of the Invention]As explained above, this invention has the following effects.

[0032](1) By using as an index the telop information given to the image, indexing to an image is automatable, and since the telop is common as compared with scenario information or title broadcast, as compared with the conventional technique, an object image is generalizable.

[0033](2) Although the amount of information of a telop as text information is small as compared with scenario information or title broadcast, develop the concept from a telop to a higher rank using a dictionary (thesaurus), and the context of an image is formed, By developing associationally by the article information which accumulated the input sentence independently of the image, and forming the input person side context, it becomes possible to detect the image associationally connected with an input sentence, and it becomes possible to compensate the amount of information with few telops.

---

[Translation done.]